INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

MRF-based Diffeomorphic Population Deformable Registration & Segmentation

Aristeidis Sotiras — Nikos Komodakis — Nikos Paragios

N° 6837

February 2008

Thème BIO

ISSN 0249-6399 ISRN INRIA/RR--6837--FR+ENG

MRF-based Diffeomorphic Population Deformable Registration & Segmentation

Aristeidis Sotiras*, Nikos Komodakis† , Nikos Paragios*

Thème BIO — Systèmes biologiques
Projet Galen

Rapport de recherche n° 6837 — February 2008 — 25 pages

Abstract: In this report, we present a novel framework to deform mutually a population of n-examples based on an optimality criterion. The optimality criterion comprises three terms, one that aims to impose local smoothness, a second that aims to minimize the individual distances between all possible pairs of images, while the last one is a global statistical measurement based on "compactness" criteria. The problem is reformulated using a discrete MRF, where the above constraints are encoded in singleton (global) and pair-wise potentials (smoothness (intra-layer costs) and pair-alignments (inter-layer costs)). Furthermore,we propose a novel grid-based deformation scheme, that guarantees the diffeomorphism of the deformation while being computationally favorable compared to standard deformation methods. Towards addressing important deformations we propose a compositional approach where the deformations are recovered through the sub-optimal solutions of successive discrete MRFs. The resulting paradigm is optimized using efficient linear programming. The proposed framework for the mutual deformation of the images is applied to the group-wise registration problem as well as to an atlas-based population segmentation problem. Both artificially generated data with known deformations and real data of medical studies were used for the validation of the method. Promising results demonstrate the potential of our method.

Key-words: mutual deformation, group-wise registration, atlas-based segmentation, Markov Random Fields

* Aristeidis Sotiras and Nikos Paragios are affiliated to Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France and to Equipe GALEN, INRIA Saclay - Île-de-France, Orsay, France
† Nikos Komodakis is affiliated to Department of Computer Science, University of Crete, Greece

Unité de recherche INRIA Futurs
Parc Club Orsay Université, ZAC des Vignes,
4, rue Jacques Monod, 91893 ORSAY Cedex (France)
Téléphone : +33 1 72 92 59 00 — Télécopie : +33 1 60 19 66 08

Recalage Difféomorphique Déformable & Segmentation de Populations d'Images basés sur des Champs Markoviens Aléatoires

Résumé : Dans ce rapport, nous présentons un cadre original pour déformer mutuellement une population de n échantillons basé sur un critère d'optimalité. Le critère d'optimalité inclut trois termes, qui visent respectivement à imposer une regularité locale, à minimiser les distances individuelles entre tous les paires d'images possibles, et enfinà intégrer une mesure statistique globale basée sur des critères de "compacité". Le problème est reformulé en tant que champ markovien, dans lequel les contraintes ci-dessus sont encodées comme des potentiels simples (global) et de potentiels par paire (régularité (coût intra-couche) et alignement par paires (coût entre-couche)). De plus, nous proposons un schéma de déformation original basé sur une grille, qui garanti l'aspect difféomorphique de la déformation tout en étant numeriquement plus efficace que les méthodes standardes. Afin de garantir une robustesse aux grande déformations nous proposons une approche composée qui retrouve les déformations itératives comme les solutions sous-optimales de champs markoviens discrets. Le paradigme résultant est optimisé en utilisant des techniques de programmation lineaire efficaces. Le cadre proposé pour la déformation mutuelle d'images est appliqué au recalage par groupes aussi bien qu'au problème de segmentation d'une population basé partir d'un atlas. Des données artificiellement crées dont les déformations sont connues ainsi que des données réeles d'études medicales etaient utilisées pour la validation de la méthode. Des résultats encourageants démontrent le potentiel de notre méthode.

Mots-clés : déformation mutuelle, recalage simultanée, segmentation à partir d'un atlas, champs markoviens aleatoires

Contents

1 Introduction

1.1 Context and Motivation

Deforming an image to match an other image or a template is a common task in a great number of problems in computer vision and medical image processing. The goal is either to establish meaningful correspondences according to a certain criterion between a pair of images [1], case of registration, or to convey information from a template, where prior information exists, to the matched image as in the case of the atlas-based segmentation problem [2]. Existing approaches to solve the fore-mentioned problems consider two images where one act as a target and the second is deformed to match the target. Such a process suffers from inherent drawbacks that stem from the fact that only the one image is permitted to deform and moreover, more importantly, from the a-priori explicit selection of the target image.

There are two major drawbacks related to the choice of the reference image. The first one is associated to the process of determining the reference. One question that would rise naturally is how well the population is represented by the pose reference. For populations that exhibit important variance, the selection of a member that is able to represent the rest in an optimal way is not a straightforward process and failure to do so can degrade the performance of the recovery of the deformations. The second major concern, when reference pose is considered, is due to the fact that its choice introduces bias to the solution as all calculations are made with the respect to it. Thus, in the case of atlas creation different reference images will result in different atlas, where each atlas is inherently biased towards the chosen reference image [3]. Such a behavior is the opposite to the one expected towards appropriate representation of the population.

The previous remarks motivate an approach where all images should be let to deform mutually while no a-priori selection of a reference pose should be made. Such an approach would permit the unbiased atlas construction that can enable the inter-member comparison and could also be used in an atlas-based population segmentation framework.

In the following, we review a number of methods that allow for the registration of a population of images.

1.2 Background and Previous Work

Several approaches that explore the mutual deformation of multiple data sets exist in the field of population registration. Population registration is defined as the identification of a homology between the input images, where the number of the images is greater than two. The major difference between all methods proposed lies in the definition of the reference frame. Significant work has been dedicated towards optimizing the selection of the reference image [4]. Once such a pose has been recovered, conventional registration methods based in pair-wise criteria can be applied [5]. The main drawback of this approach is that reference must be a member of the population. A natural way, to address this shortcoming is to perform simultaneously the registration of all subjects to a common pose that is not predefined.

Towards this end, in [6] all the pair-wise registrations between members of the population were considered and a mean model was created by composing the mean deformations for each member into a mean deformation. This improves the performance with respect to the pair-wise registration but still suffers from not taking into account the global statistics of the population.

The idea of using the statistics of the population to construct a template based on which the similarity criterion is a natural extension of the above mentioned methods. In [7], both the deformation and the template selection are optimized at the same time and a squared error dissimilarity measure was used. In [8] a normalized mutual information criterion in a Free Form Deformation framework was used. The template was defined as the average transformation of the population. It should be noted that in approaches as the ones presented in [9, 10] a minimum description length criterion was used to perform group-wise registration and model building. The main limitation of these methods is relevant to the "use" of a template which could greatly influence the obtained results. For example, the use of such methods becomes less evident when a important number of diseased patients are part of the population, or the population involves different modalities. In [11], the idea of performing template-free group-wise registration was proposed where the objective criterion used is the sum of univariate entropies along pixel stacks. This method was applied in the problem of group-wise registration in [12] by combining affine transforms and a stochastic gradient descent optimization framework. This affine congealing framework was further extended by [13] to include Free Form Deformations.

State of the art population-registration methods have either opted for pair-wise registrations and template construction or global statistical measurements. Both components exhibit strengths and limitations. In both cases, modularity and computational performance are the main limitations. The main difficulty associated with group-wise registration refers to the high computational demand, memory requirements and running times in particular when addressing the deformable case. These constraints were partially addressed through efficient computational approximations [12, 14] but still when referring to the deformable case, the spectrum of these methods is far from being satisfactory.

However, in terms of modularity with respect to the selected registration criterion most of the proposed methods fail short when a different deformation model or registration criterion (pair-wise or global) is to be considered. One would like for example, if the population involves different modalities, to use similarity metrics that do change according to the nature of the sample. Furthermore, the ability to encode various global statistical measurements towards evaluating the statistical coherence of the entire population would be of great use. Last but not least, the flexibility in terms of the nature of transformation should be considered.

1.3 Contributions

In this report, several issues are addressed. The mutual deformation of a population of images is formulated efficiently with the use of a graphical model approach [Fig. 1]. The latent variables of the model are n-deformations (Hermite-based polynomials) of the popu-

lation examples and the "optimal" reference pose. The pose variables are connected with the observations and the corresponding deformation variables towards measuring the statistical compactness of the registration result at the pixel level. The registration variables are inter-connected and aim to decrease the cost of pair-wise comparisons between individual examples. Last, but not least the registration variables within an image are connected so as to impose smoothness. The resulting paradigm can easily encode different deformation interpolation methods, local similarity metrics and global statistical measurements while being computational efficient [when compared with the state of the art methods]. This graphical model is expressed in the form of a MRF.

Various grid-based deformations were considered while a Hermite-based deformation scheme is introduced. The proposed deformation method is a computationally efficient variant to the commonly used grid-based deformations, while at the same time exhibits certain desirable properties such as preserving the diffeomorphism of the deformation under certain conditions.

The mutual deformation of a population of images framework is applied to the atlas-based segmentation problem and the notion of mutual population segmentation is introduced. The aim is to combine prior knowledge along with consistency through the simultaneous segmentation of the whole population. This is achieved through the mutual deformation of the population members towards the atlas, while at the same time being constrained through a simultaneous all-to-all deformable diffeomorphic registration. The latter will impose consistency with respect to the population segmentation results.

The efficiency of the proposed methodology was tested for both the problems of population registration and the atlas-based mutual population segmentation. For the first problem, experiments were carried out using artificially generated data whose deformations are known as well as real data from MRI acquisitions of the calf muscles and chest radiographs. As far as the second problem is concerned, results are reported for a publicly available data set of posterior-anterior chest radiographs.

The remainder of this report is organized as follows: In section 2 the concept of our diffeomorphic mutual population deformation based on local and global criteria approach is described while its MRF variant is introduced in section 3. In section 4 the experimental validation is presented, followed in section 5 by the discussion that concludes this report.

2 Global and Local Criteria for Mutual Population Deformation

In order to introduce the concept of our approach let us consider n images $\{I_1, ..., I_n\}$. We can assume that its image is described by intensity values $I_i(\mathbf{x}_i)$ for different image domains $\Omega_i, \mathbf{x}_i \in \Omega_i$. Without loss of generality a common reference frame can be defined (one should note that this assumption refers to the transformation domain definition and not to an image template). The aim of the mutual population deformation is to determine a transformation $\mathbf{T}$ which maps mutually points from the image space, defined by all n images, to points to

the reference frame Ω_R

$$\mathbf{T} = \{T_i : \mathbf{x}_R = T_i(\mathbf{x}_i), i = 1, ..., n\}. \quad (1)$$

such as an optimality criterion is satisfied.

2.1 Deformation Model

Let us consider a grid-based deformation model mostly to facilitate the introduction of the method and maintain the ability to encode different interpolation methods. Furthermore, we assume that the transformation is one-to-one and invertible. The deformation of an object is achieved by manipulating an underlying mesh of control points; the resulting deformation controls the shape of the object. We superimpose a deformation grid $G_i : [1, K] \times [1, L]$ (usually K and L are smaller than the dimensions of the image) onto each one of the images $I_i, i = 1, ..., n$. The central idea of our approach is to deform the grids simultaneously (with a 2D displacement vector $\mathbf{d}_{\mathbf{p}_i^k}$ for each control point k belonging to the grid G_i) such that the optimality criterion is optimized. In this case, the transformation of an image pixel $\mathbf{x}_i = (x_i, y_i) \in \Omega_i$ can be written as

$$T_i(\mathbf{x}_i) = \mathbf{x}_i + D_i(\mathbf{x}_i) \quad (2)$$

where $D_i(\mathbf{x}_i)$ is defined as:

$$D_i(\mathbf{x}_i) = \sum_{\mathbf{p}_i^k \in G_i} \eta(|\mathbf{x}_i - \mathbf{p}_i^k|)\mathbf{d}_{\mathbf{p_i^k}}. \quad (3)$$

Without loss of generality, in such a theoretical setting, we can consider the cubic-B splines Free Form Deformations. In this case the former equation can be written as:

$$D_i(\mathbf{x}_i) = \sum_{r=0}^{3}\sum_{c=0}^{3} B_r(u)B_c(v)\mathbf{d}_{\mathbf{p}_i^{u_0+r,v_0+c}}. \quad (4)$$

In the previous equation, $u_0 = \lfloor x_i/\delta_x \rfloor, v_0 = \lfloor y_i/\delta_y \rfloor, u = x_i/\delta_x - \lfloor x_i/\delta_x \rfloor$ and $v = y_i/\delta_y - \lfloor y_i/\delta_y \rfloor$. B_r represents the rth basis function of the B-spline and $\delta_x = \frac{S_x}{K-1}, \delta_y = \frac{S_y}{L-1}$, (where S_x, S_y are the dimensions of the image along x- and y-axis respectively), denotes the control point spacing.

In the same theoretical setting, in order to render the computations of the deformation field even more efficient and the same time localize the influence of the control points while at the same time being smooth enough, we considered the Hermite splines [15]. In this case equation 3 takes the following form:

$$D_i(\mathbf{x}_i) = \sum_{r=0}^{1}\sum_{c=0}^{1} H_r(u)H_c(v)\mathbf{d}_{\mathbf{p}_i^{u_0+r,v_0+c}} \quad (5)$$

where H_l represents the lth basis function of the Hermite spline and the rest are defined as previous.

Given the above deformation model, the population of images will be deformed by performing global statistical compactness measurements and pair-wise comparisons.

2.2 Population-wise Global Comparisons

The first term of the objective criterion to be minimized is the global statistical one. Let $\pi(\mathbf{i}(\mathbf{x}))$ be the distribution of the corresponding sample intensities at images $\{I_1, \cdots, I_n\}$, or $\mathbf{i}(\mathbf{x}) = \{I_1(T_i^{-1}(\mathbf{x})), \cdots, I_n(T_n^{-1}(\mathbf{x}))\}$. In statistics, one can associate a random variable to a measure of compactness with respect to this density. Examples can refer to standard deviation, higher order moments, entropic measurements, etc. We introduce the following global measurement towards population registration

$$E_{global}(T_1, \cdots, T_n) = \iint_{\Omega_R} \gamma(\pi(\mathbf{x}))d\mathbf{x} \tag{6}$$

with γ being a monotonic function inversely proportional to the compactness of the distribution at $\mathbf{x}$. For example, we can consider the entropy is estimated by using the histograms of the intensity values. The justification of using this approach is that as the images are mutually deformed to be aligned properly the compactness of the probability distribution should increase and the intensity values at corresponding coordinate locations from all the images will form a low entropy distribution. Such an objective function introduces the inverse transformation, that is challenging from theoretical and practical point of view when referring to deformable deformation. An alternative criterion that can be considered is using the forward transformations and measure the similarity of the images on the intersection of the deformed images, or

$$E_{global}(T_1, \cdots, T_n) = \int \cdots \int_{\Omega_i \cup \cdots \cup \Omega_n} \phi(T_1(\mathbf{x}_1), \cdots, T_n(\mathbf{x}_n))\gamma(\lambda(\mathbf{x}_1, \cdots \mathbf{x}_n))d\mathbf{x}_1 \cdots d\mathbf{x}_n \tag{7}$$

where $\lambda(\mathbf{x}_1, \cdots \mathbf{x}_n) = \pi(I_1(T_1(\mathbf{x}_1)), \cdots, I_n(T_n(\mathbf{x}_n)))$ and ϕ is a dirac-driven function whose role is to define which pixels correspond to the same position at the reference pose defined as follows: $\prod_{(i,j)\in[1,n]\times[1,n]} \delta_\alpha(|\mathbf{x}_i - \mathbf{x}_j|)$. We can rewrite the above objective function using the notion of control points, where each pixel is back-projected to one of the points of the grid, or

$$E_{global}(T_1, \cdots, T_n) = \sum_{\mathbf{p}_j^k} \int \cdots \int_{\Omega_i \cup \cdots \cup \Omega_n} \phi(T_1(\mathbf{x}_1), \cdots, T_n(\mathbf{x}_n))\eta_s^{-1}(\mathbf{p}_j^k)\gamma(\lambda(\mathbf{x}_1, \cdots \mathbf{x}_n))d\mathbf{x}_1 \cdots d\mathbf{x}_n \tag{8}$$

where we have summed over all the nodes k belonging to all grids j. The back-projection function is defined as,

$$\eta_s^{-1}(|\mathbf{x}_i - \mathbf{p}_i|) = \frac{\eta(|\mathbf{x}_i - \mathbf{p}_i|)}{\int_{\Omega_i} \eta(|\mathbf{y}_i - \mathbf{p}_i|)d\mathbf{y}_i}. \tag{9}$$

Examples of such global measurement can be the standard deviation, the skewness, the kurtosis, the Shannon entropy as considered in the case of congealing, or more advanced statistical compactness measurements.

2.3 Pair-wise Local Comparisons

Let $\rho_{i,j}(\cdot)$ be a similarity measurement used to compare the visual information for the images i and j. Then, if (without loss of generality) we consider for example pixel-based measurements, the pair of deformations T_i, T_j, should minimize the distance between the two individual images in the reference domain:

$$E_{local}(T_i, T_j) = \iint_{\Omega_R} \rho_{i,j}(I_i(T_i^{-1})(\mathbf{x}), I_i(T_j^{-1}(\mathbf{x})))d\mathbf{x} \tag{10}$$

where T_i^{-1}, T_j^{-1} are the inverse transformations between the reference domain and the domains Ω_i, Ω_j and Ω_R is the common reference domain. This criterion simply measures the similarity for every pixel of the reference pose with respect to the corresponding observations at the origin pixels at $I_i(\cdot)$, and $I_j(\cdot)$. Similarly, to the case of the global statistical measurements, we introduce the following local measurement towards population registration

$$E_{local}(T_i, T_j) = \iint_{\Omega_i \cup \Omega_j} \phi(|T_i(\mathbf{x}_i) - T_j(\mathbf{x}_j)|)\rho_{i,j}(I_i(T_i(\mathbf{x}_i)), I_j(T_j(\mathbf{x}_j)))d\mathbf{x}_i d\mathbf{x}_j \tag{11}$$

In simple words, this quantity evaluates the pertinence of the correspondences between the two images using both definition domains Ω_i, Ω_j where only the pixels for which correspondences between the two images have been found are considered. It is the role of the $\phi(\cdot)$ dirac-type function to which determine pixels $\mathbf{x}_i \in \Omega_i$ and $\mathbf{x}_j \in \Omega_j$ correspond to the same pixel $\mathbf{x} \in \Omega_R$ where the comparison takes place. We can rewrite the above objective function using the notion of control points, where each pixel is back-projected to one of the points of the grid, or

$$E_{local}(T_i, T_j) = \sum_{k \in G_i} \sum_{m \in G_j} \iint_{\Omega_i \cup \Omega_j} \bar{\phi}(\mathbf{x}_i, \mathbf{x}_j)\eta_p^{-1}(|\mathbf{x}_i - \mathbf{p}_i^k|, |\mathbf{y}_j - \mathbf{q}_j^m|)\bar{\rho}_{i,j}(\mathbf{x}_i, \mathbf{x}_j)d\mathbf{x}_i d\mathbf{x}_j \tag{12}$$

where $\bar{\rho}_{i,j}(\mathbf{x}_i, \mathbf{x}_j) = \rho_{i,j}(I_i(T_i(\mathbf{x}_i)), I_j(T_j(\mathbf{x}_j)))$ and $\bar{\phi}(\mathbf{x}_i, \mathbf{x}_j) = \phi(|T_i(\mathbf{x}_i) - T_j(\mathbf{x}_j)|)$. The weighting function η^{-1} computes the influence of the image point $\mathbf{x}_i, \mathbf{x}_j$ to the control points $\mathbf{p}_i, \mathbf{q}_j$. For example, if we consider the case of closest-neighbor interpolation, then a given pixel in the image will only contribute to the closest control point with a coefficient equal to one.

The inverse function takes the following form:

$$\begin{aligned} \eta_p^{-1}(\mathbf{p}_i, \mathbf{q}_j) &= \eta_p^{-1}(|\mathbf{x}_i - \mathbf{p}_i|, |\mathbf{x}_j - \mathbf{q}_j|) \\ &= \frac{\eta(|\mathbf{x}_i - \mathbf{p}_i|)\eta(|\mathbf{x}_j - \mathbf{q}_j|)}{\int_{\Omega_i \cup \Omega_j} \delta_\alpha(|T_i(\mathbf{y}_i) - T_j(\mathbf{y}_j)|)\eta(|\mathbf{y}_i - \mathbf{p}_i|)\eta(|\mathbf{y}_j - \mathbf{q}_j|)d\mathbf{y}_i d\mathbf{y}_j}. \end{aligned} \tag{13}$$

In that way we are able to consider the influence of the image point $\mathbf{x} \in \Omega_R$ to the nodes of the grids that are superimposed onto the two images. This inverse function will be used during the evaluation of the pair-wise potentials as it should be clarified in the continuation.

This criterion can be extended to deal with the case of n-images by simply considering all possible pairs of images, or :

$$E(T_1, \cdots, T_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} E_{local}(T_i, T_j) \tag{14}$$

This term will introduce consistent pair-wise deformations but will not introduce compactness in terms of statistical behavior for the observations once brought to the reference pose.

2.4 Smoothness Constraints

Medical images capture properties of spatially continuous anatomical structures, therefore it is natural to assume that the deformation applied to them should be locally smooth. Opposite to the former cases, this constraint should applied to each grid separately. We can either define it on the image or the transformation grid (the two definitions become equivalent with back-projection from the image to the grid). For the shake of simplicity we define it directly on the deformation of the grid, or

$$E_{global}(T_1, \cdots, T_n) = \sum_{i=1}^{n} \sum_{\mathbf{p}_i^k \in G_i} \psi(|\nabla_{G_i} \mathbf{d}_{\mathbf{p}_i^k}|) \tag{15}$$

where ψ is a convex function imposing smoothness.

The optimal parameters of the deformation should be determined by an energy minimization step. Gradient descent method is the most common approach to address the minimization task [16], but suffers from important limitations. Among its limitations are the inability to guarantee the recovery of the global minimum and its computational inefficiency. Moreover, it is not modular since the derivative of the energy depends both on the model and the criterion used. The use of an alternative optimization strategy that will be able to approach the optimal solution with a reasonable computational effort and at the same time encompass different optimization criteria is therefore necessary. Strategies that comply with the proposed characteristics that the optimization scheme should exhibit, can be found in the area of discrete optimization.

3 Mutual Population Deformation: Discrete Domain

Strategies that comply with the proposed characteristics that the optimization scheme should exhibit, can be found in the area of discrete optimization. In order to able to use discrete optimization schemes the deformation space should be quantized. Let $\Theta = \{\mathbf{d}^1, .., \mathbf{d}^q\}$ be a quantized version of the deformation field, then a discrete set of labels $L = \{l^1, ..., l^q\}$ can be corresponded to it. A label assignment $l^\xi_{\mathbf{p}}$, where $\xi \in \{1, \cdots, q\}$, to a grid node $\mathbf{p}$ is associated with displacing the node by the corresponding vector $\mathbf{d}^{l^\xi_{\mathbf{P}}}$. If a label is assigned to every node we get a discrete labeling $\mathbf{l}$. The displacement field associated with a certain

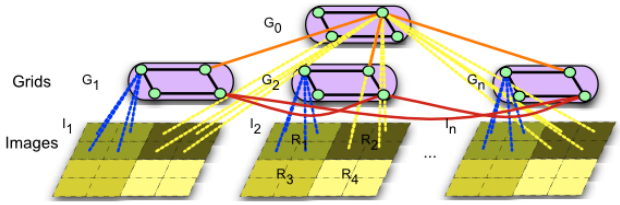


Figure 1: The node and the edge system of the constructed graph. With blue color the relationship between the grid nodes and the images is depicted (deformation model). The black edges represent the smoothness terms while the red ones encode the local dissimilarity measure. The global relationship between all the nodes at respective places in the grids is shown by the yellow edges. (For clarity a fraction of the edges is shown.)

labeling $\mathbf{l}$ becomes

$$D(\mathbf{x}) = \sum_{\mathbf{p}\in G} \eta(|\mathbf{x}-\mathbf{p}|)\mathbf{d}^{l_\mathbf{p}}. \tag{16}$$

By applying this quantization of the deformation space one would like to reformulate the problem as a discrete multi-labeling problem. A common model for representing such problems are Graphical Models and MRFs. In the context of population registration, the graphical model will involve three terms, one singleton that measures the compactness and two pair-wise, one that account for smoothness at each deformation field and one that enforces pair-wise correspondences.

$$\begin{aligned} E_{GM}(G_0, T_1\circ G_1, \cdots, T_n \circ G_n,) = \sum_{i=0}^{n}\sum_{\mathbf{p}\in G_0} V_\mathbf{p}(l_\mathbf{p}) + \\ \sum_{i=0}^{n}\sum_{\mathbf{p}\in G_i}\sum_{\mathbf{q}\in(N(\mathbf{p})\cap G_i)} V_{\mathbf{pq}}(l_\mathbf{p}, l_\mathbf{q}) + \sum_{i=0}^{n}\sum_{\mathbf{p}\in G_i}\sum_{\mathbf{q}\in(N(\mathbf{p})\setminus G_i)} V_{\mathbf{pq}}(l_\mathbf{p}, l_\mathbf{q}) \end{aligned} \tag{17}$$

where $V_\mathbf{p}(\cdot)$ are the unary potentials, $V_{\mathbf{pq}}(\cdot,\cdot)$ are the pair-wise potentials and N represents the neighborhood system of the nodes [Fig:1].

The main challenge of discrete optimization methods is the quantization of the search space since it seeks for a compromise between computational complexity and the ability to capture a good minimum. Dense quantization of the solution space in combination with one shot minimization strategies can approach a good minimum at the cost of a high computational demand that renders the problem intractable. On the other hand, coarse quantization of the deformation field result in an efficient optimization scheme but the ability to approach a good minimum is degraded. A good compromise is achieved through a compositional approach, where the final solution is obtained through successive optimization problems with respect to the deformation increment towards minimizing the objective function [17, 1]. Thus, by keeping the set of the labels in a reasonable size it becomes possible to approximate the optimal solution in an efficient way.

3.1 Mapping of the Objective Function to the Graphical Model

Mapping global, local and smoothness costs to the graphical model consists of converting them to singleton and pair-wise terms. The most challenging case is the global cost due to the fact that in order to be properly determined it requires higher order cliques, while our model consists of cliques of order two. The mapping of the other two terms is straightforward. It is possible to introduce high order cliques but such an energy cannot be (at least easily) minimized [18], thus the global term will be approximated.

We consider an approximation of the global cost that consists of assuming that for a given node $\mathbf{p}$ of a given deformation field/image i, the rest of the images do not move within the current iteration. This assumption is considered for all nodes, and for all deformation fields within a given iteration and therefore is not restrictive and quite common in minimizing graphical models through expansion moves. Then, the cost of a deformation will depend only on the label of this node, or,

$$V^t_{\mathbf{p_i^k}}(l_{\mathbf{p_i^k}}) \approx \int \cdots \int_{\Omega_1 \cup \cdots \cup \Omega_n} \eta_s^{-1}(\mathbf{x}, \mathbf{p}_i^k)\phi(T_1^{t-1}(\mathbf{x}_1), \cdots, T_i^t(\mathbf{x}_i), \cdots T_n^{t-1}(\mathbf{x}_n)) \\ \gamma(\lambda(I_1(T_1^{t-1}(\mathbf{x}_1)), \cdots, I_i(T_i^t(\mathbf{x}_i)))d\mathbf{x}_1 \cdots d\mathbf{x}_n$$

where

$$\eta_s^{-1}(|\mathbf{x}_i - \mathbf{p}_i|) = \frac{\eta(|\mathbf{x}_i - \mathbf{p}_i|)}{\int_{\Omega_i} \eta(|\mathbf{y}_i - \mathbf{p}_i|)d\mathbf{y}_i}. \tag{18}$$

In our approach we have considered two global statistical measurements, a congealing-like global cost that aims at minimizing the entropy of the pixel distributions upon registration and one that aim to minimize the pixel-wise standard deviation. Therefore this term corresponds to the G_0 graphical model variables.

As far as the pair-wise potentials are concerned, two different cases have to be discerned. First of all, there are the pair-wise potentials that encode the local smoothness constraint of the deformation field which will be referred as intra pair-wise potentials. A simple smoothness term can be defined as a distance function computing the magnitude of vector differences

$$V_{\mathbf{p_i q_i}}(l^\xi_{\mathbf{p}_i}, l^\nu_{\mathbf{q}_i}) = \beta_{intra}|\mathbf{d}^{l^\xi_{\mathbf{p}_i}} - \mathbf{d}^{l^\nu_{\mathbf{q}_i}}| \tag{19}$$

where β_{intra} is a weighting that it has been considered to be the same over the spatial domain and $d^{l^\xi_{\mathbf{p}_i}}$ actually is the displacement that corresponds to assigning a label l^ξ to the node $\mathbf{p}_i$.

The second type of pair-wise potential that will be referred as inter pair-wise potential encodes the distance between the pair of images. A point to be cleared out before defining the inter pair-wise potentials is that this type of pair-wise potential is only calculated between the nodes $\mathbf{p}_i^k, \mathbf{q}_j^k$. These nodes are placed in respective places k in grids that belong to two different images i and j. The inter pair-wise potential are defined as

$$V_{\mathbf{p}_i^k \mathbf{q}_j^k}(l_{\mathbf{p}_i^k}, l_{\mathbf{q}_j^k}) \approx \int_{\Omega_i \cup \Omega_j} \phi(|T_i(\mathbf{x}_i) - T_j(\mathbf{x}_j)|)\eta_p^{-1}(\mathbf{p}_i^k, \mathbf{q}_j^k) \cdot \rho(I_i(T_i(\mathbf{x}_i)), I_j(T_j(\mathbf{x}_j)))d\mathbf{x}_i d\mathbf{x}_j \tag{20}$$

where ρ is a dissimilarity measure (sum of square differences as well as sum of absolute differences have been tested). The inter pair-wise potentials are multiplied by a factor β_{inter}, which normally is different than β_{intra}.

To minimize the successive MRFs, that is to assign a label l to all the nodes $\mathbf{p}$ of the constructed graph, a state of the art minimization method will be used. The method is called fast-PD and is detailed in [19].

The last constraint to be addressed refers to the diifeomorphic property of the proposed population registration framework. This can easily introduced by imposing hard constraints to the allowed deformations [20]. As the solution space is actually a quantized version of the deformation space and corresponds to the label set, it is simple to restrict the value that the maximum displacement can take in terms of the grid spacing. Towards imposing the diffeomorphic property on the deformation field, the maximum displacement is restricted to be $0:4$ times the grid spacing [21] in the case of the cubic-B splines. Following [21], the bound for the maximum displacement, in the case of the cubic Hermite spline, is proven to be $0:25$ times the grid spacing.

4 Experimental Validation

4.1 Group-wise Registration

A natural way to apply the proposed framework is for the solution of the group-wise registration problem where the members of the population are mutually deformed to reach the same pose based on the minimization of an optimality criterion. The proposed method was tested by using three data sets, an artificial and two real ones, to validate its capability to register simultaneously a group of images.

4.1.1 Data

A synthetic brain data set was created by applying simulated deformations. The warping is done by using cubic B-spline FFD model with two different levels of deformation grids. The maximum displacement applied was constrained to be -14.6 to 14.6 millimeters in both directions in order to preserve the diffeomorphism of the transformation. In order to better assess the quality of the registration, landmarks were placed on the original image so as to permit us to evaluate visually the convergence to a common frame as the landmarks approach each other.

The second data set consists of 18 MRI T1 images of the calf, and manual expert segmentations of five major muscle groups. The images were acquired with a 1.5 T Siemens scanner, with parameters TR=711 /TE=11. Each volume consists of 90 slices of 4mm thickness with voxel spacing of $0.7812 \times 0.7812 \times 4$ mm. It is important to note that the data set consists of members that exhibit great variability. Images that belong to patients as well as ones of healthy subjects are part of the data set 2.

The third data set consists of chest radiographs that are taken from the JSRT database [22]. This database contains 247 PA chest radiographs collected from 13 institutions in Japan and one in the United States. The images were scanned from films to a size of 2048×2048 pixels, a spatial resolution of 0.175 mm/pixel and 12 bit gray levels. 154 images contain exactly one pulmonary lung nodule each; the other 93 images contain no lung nodules. The images where segmented and made publicly available by the authors of [23]. From this publicly available data base only the 93 images that contain the no lung nodules were used. The images were scaled down to 256×256 and no further processing took place. From this data set 20 images were selected randomly in order to test the performance of the proposed algorithm. A sample of the dataset finally used is depicted in Fig.3, where both the images and their lung field segmentations are presented.

4.1.2 Implementation Details

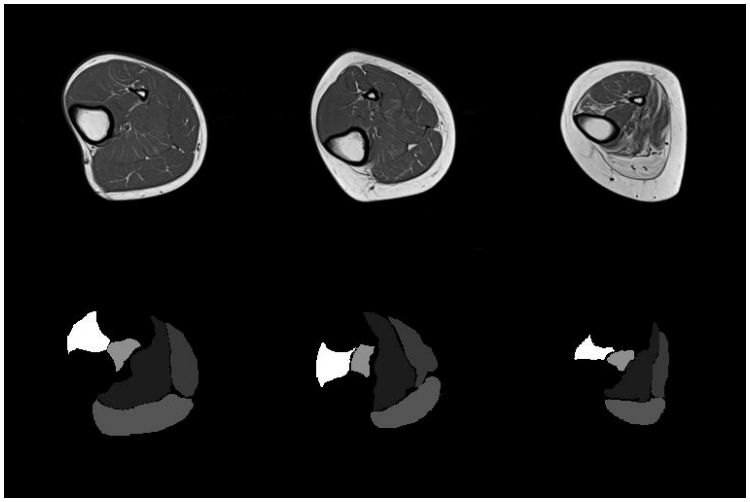


Figure 2: Typical images of the muscle data set. In the first row, T1 images of the calf for 3 patients are presented. In the second row, the respective segmentations for five major muscle groups are given.

The sum of absolute differences was considered as pair-wise cost function between the two deformed images. For the first two data sets the entropy of the aligned images per pixel was used to impose global statistical compactness, while for the third one the pixel-wise standard deviation was considered. Regarding the parameters of the method, α was set to 10, β_{inter} was set to 1 and β_{intra} to 0.1. We have used a multiscale implementation with 3 levels, an initial grid resolution of 8×8, and a final one of 32×32. A number of $2 \times 8 + 1$ labels were used per iteration cycle, sampled along the principal horizontal and vertical directions. The Hermite-splines based deformation scheme was used for these experiments.

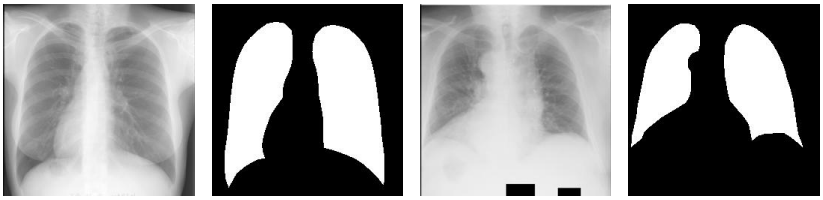


Figure 3: Typical images of the dataset: (a) Radiograph 1, (b) Radiograph 2, (c) Segmentation of the lung fields for the first image, (d) Segmentation of the lung fields for the second image.

4.1.3 Results

The qualitative results of the group-wise registration of the muscle data are presented in [Fig4]. Comparing visually the mean and the variance image of the population before and after the group-wise registration the success of the registration process can be assessed qualitatively. The mean image is far more sharp than the one before the registration process, while the variance image emphasizes the decrease of the intensity differences along the registered data. To further appraise the performance of the proposed method, it was compared to a state of the art pair-wise registration method [1]. Similar parameters and deformation grids were used for both methods with the difference that for the group-wise registration scheme hermite weighting functions were used in the place of cubic B-splines. The performance of the pairwise registration was exhaustively evaluated as all possible images were used as targets. The distributions of the dice values for each image target are reported in [Fig.5], where a boxplot is given for every image target. The results for the pair-wise registration are given from column 1 to 18 while the last column corresponds to the results obtained by the proposed group-wise registration framework. By simple observation of the graph it can be concluded that the median value of the Dice coefficient attained through the group-wise registration outperforms the pair-wise method for the majority of the target images.

For the case of the radiograms we can remark that the mean segmentation gets sharper. Similarly to the muscle data set, quantitative results are presented for the case of radiographs in [Fig 7]. In columns 1 to 20 the dice distributions obtained by performing pair-wise registrations are reported. In the last column the results obtained through pairwise registration are shown. By simple observation of the graph, it can be concluded that the group-wise registration method performs better than the pair-wise one. In order to draw conclusion in a more objective way, a paired statistical Student t-test with a significance level of 0.05was performed for comparison. The results of the test prove that the group-wise method performed better in 12 cases, while the pairwise one performed better for 6 cases.

The quantitative results presented for both the muscle and radiographs data set suggest that considering the population as a whole and registering subjects jointly brings the population into better alignment than matching each subject to a target image. This is implied by the decrease of the dispersion of the Dice values that is observed in the group-wise case.

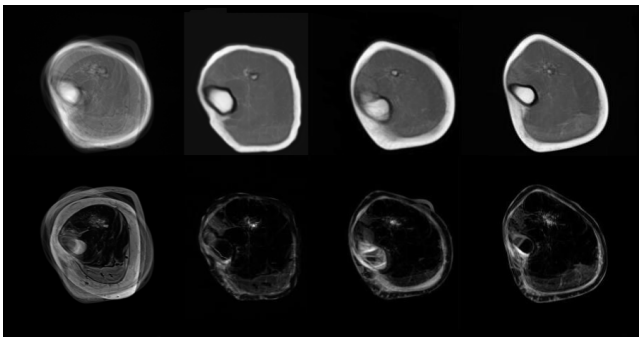


Figure 4: Results obtained for the muscle image data set. In the first row the mean image is presented, while in the second the variance. From left to right, the initial images, the result of the groupwise registration, the result of two pair-wise registrations.

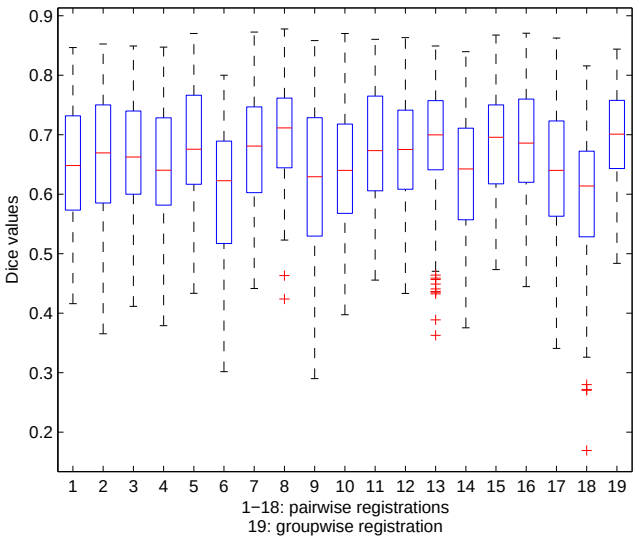


Figure 5: Comparison between group-wise and pair-wise registration for the muscle data set. The DICE with respect to the plausible individual template choices are compared with the population registration result.

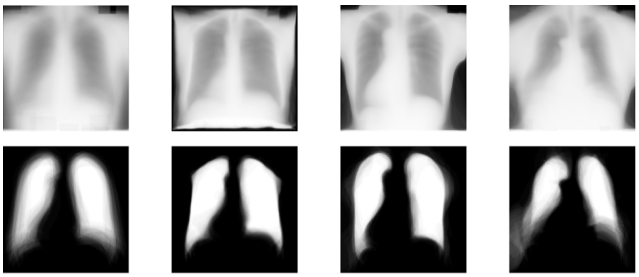


Figure 6: Results obtained for the radiographs data set. In the first row the mean grey scale image is presented, while in the second the mean segmentation image. From left to right, the initial images, the result of the groupwise registration, the result of two pair-wise registrations.

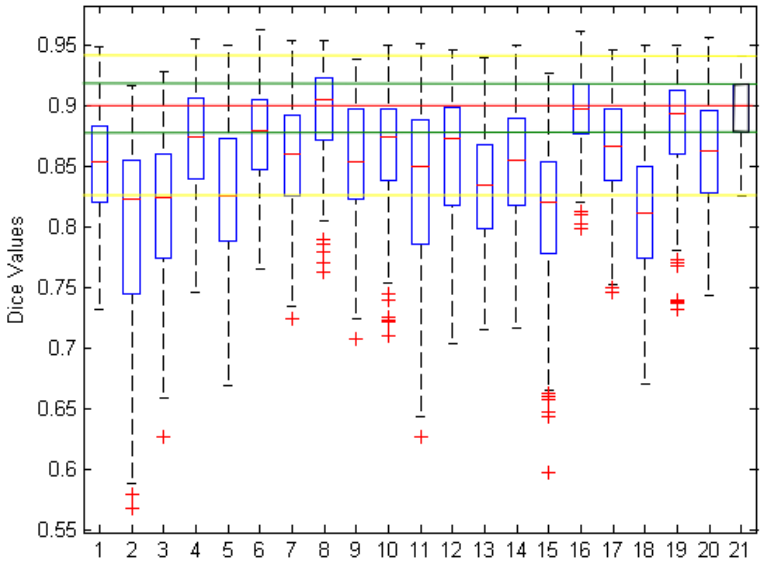


Figure 7: Comparison between group-wise and pair-wise registration for the chest radiograms images. The DICE with respect to the plausible individual template choices are compared with the population registration result.

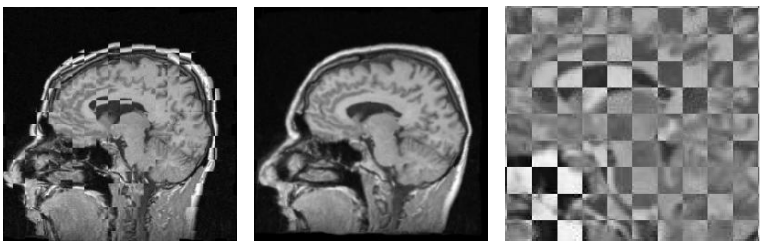
The quantitative results presented in the figures [Fig. 5, 7] also point out the intrinsic drawbacks of the pair-wise registration process whose performance is greatly influenced by the choice of the target image.

The qualitative results presented for the brain data set point out the merits of the proposed method. As it is depicted in Fig.9 the proposed method manages to bring the whole population to the same pose. This is made evident as the mean image of the population becomes sharper but more importantly as the respective landmarks in all images converge to the same positions. More qualitative results on this data set are presented in Fig.8, where a checkerboard is created from both the input and output images. The fact that the tiles of the input images that compose the checkerboard can be easily seen implies that the images exhibit great variations. As it is evident from the smooth checkerboard of the output images, the algorithm has managed to bring all members to the same pose.

The running time of our approach using Matlab implementation for a population registration of 20 examples (256×256), and a final resolution grid of 32×32 per image is approximately 30 min on an Apple Mac with 4GB memory and 2.5GHZ Processor. However, since our graph has similar complexity to the one reported in [1] and the same optimization technique is used, a C++ implementation should bring this running time down to a couple of minutes.

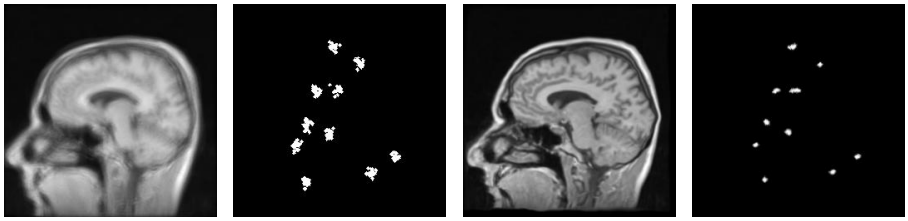
4.2 Atlas-based Mutual Population Segmentation

Another field where the efficiency of such an approach could be evaluated is the one of atlas-based segmentation. Atlas-based methods [2] and segmentation through registration is a popular research field. The basic idea is to deform an "average" organ representation towards the image and use it to obtain the solution through the deformation of the atlas



(a) (b) (c)

Figure 8: Visual results for the brain dataset: (a) Checkerboard of input images, (b) Checkerboard of output images, (c) Detail of the checkerboard of the output images. Note, that in the latter case, dark areas should be aligned with bright areas.



(a) (b) (c) (d)

Figure 9: Visual results for the brain dataset: (a) Mean image before the registration, (b) Image presenting the positions of the landmarks before the registration, (c) and(d) respective images after the registration.

segmentation map. Such a concept has numerous advantages in particular when organ delineation is not straightforward. One can cite for example lack of edges (radiographic images) or presence of quite complex texture and similarities with the rest of the region being imaged (different classes of the human skeletal muscle). In such a context the registration can be either global (rigid, affine, etc.) or local (deformable) and results will vary and being quite sensitive from one example to another according to the performance of the registration method.

Such a process - despite being quite efficient in numerous applications - does not take into account the entire set of examples to be segmented and treats individually each case. One can expect that a population-driven approach will perform better. The central idea is to simultaneously segment all examples while imposing image-based consistencies between them. Such a concept is well motivated from a number of recent studies in the field of population registration [24, 12, 8] which demonstrate that pair-wise methods are quite biased and their performance is significant lower to the population ones.

In this report, we introduce the concept of population segmentation through atlas and population deformable registration. Our approach aims to register the atlas to the individual examples, while at the same time register all examples among them. The later, will impose consistency with respect to the population segmentation results. In both cases, we seek for deformable deformations and all of them are simultaneously recovered.

The algorithm was validated for segmentation of the lung fields in a population of standard posterior-anterior chest radiographs of a publicly available database.

4.2.1 Data

From the previously presented chest radiographs data set, 20 images where randomly selected to create the atlas and another 20 of them were used to test the performance of the proposed method. By simple observation of the images 3, it becomes evident that the images contain no great edge content while they exhibit important intensity variability. What should be

pointed out though, is that the form of the lung fields also exhibit great variability thus making the problem of the segmentation a challenging one.

4.2.2 Implementation Details

The compositional approach was combined with a multiscale grid implementation and a Gaussian pyramid representation of the images. Thus, the first iterations are performed by using the coarser grid and the coarser level of the Gaussian pyramid which are then refined during the following optimization cycles. A number of $2 \times 8 + 1$ labels were used per iteration cycle, sampled along the principal horizontal and vertical directions. Regarding the parameters f the method, this time only the local pair-wise comparisons were used, thus α was set equal to 0 while β_{inter} to 0.8 and β_{intra} to 0.1. The pair-wise similarity criterion that was used was the Sum of Absolute Differences while the cubic-B splines Free Form Deformations were considered.

4.2.3 Results

The Dice similarity coefficient, the sensitivity and the specificity were computed in order to evaluate the segmentation results. The obtained quantitative measurements are graphically depicted in the form of a boxplot in Fig.10. As it can be easily noted, the proposed method manages to segment the lung fields with success as the dice coefficients imply. To be able to visually assess the performance of the method, the comparison between the estimated solution and the ground truth in the common pose is presented in Fig.11. Qualitatively, it can be argued that there is a good accordance between the two borders. However, inaccuracies are present, mainly in zones where ambiguity concerning the border of the object exists and especially for images that exhibit great deformations. Tackling the current limitations is subject of ongoing research.

5 Discussion

We have presented a new method that is based on discrete optimization techniques and that combines a global entropic criterion along with standard pair-wise dissimilarity measures to perform mutual population deformation. Furthermore, we have introduced the novel concept of atlas-based mutual population segmentation method where segmentations do exchange information and convergence/simultaneous labeling is reached once a consensus has been found among them. One important characteristic of the proposed approach is its computational efciency that it is enhanced by the use of state of the art MRF optimization techniques as well as by the use of a novel efcient diffeomophic grid based deformation model. A key characteristic of the proposed approach is its versatility. Not only is it able to encompass different similarity measures and smoothness functions, but also different approaches can be modeled through our proposed framework. As the results presented show, our method performs well both on synthetic and real medical data in both group-wise and atlas-based mutual population segmentation framework.

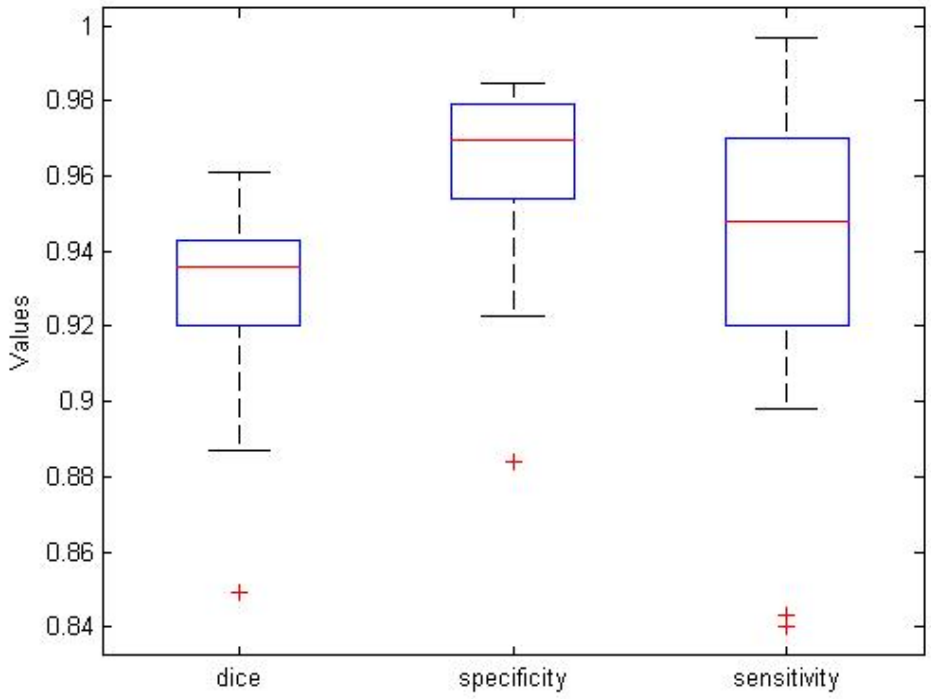


Figure 10: Graphical representation of the measured Dice coefficient, the specificity and the sensitivity.

We believe that this is a promising direction of research and thus in the immediate future, we are going to test it on 3D data in combination with a greater range of dissimilarity measures. For example, introducing more complex representations with respect to the atlas through more appropriate statistical characterization of the training data is a straightforward extension. One second direction that should be investigated is the use of non uniform deformation grids that will be more dense in areas that exhibit larger deformations. The use of prior geometric knowledge with respect to the set of admissible deformations is also a promising direction. Finally, a direction that should be explored is the possibility to relax the connectivity of the graph created, as it is going to enable us to apply our method in a bigger set of data.

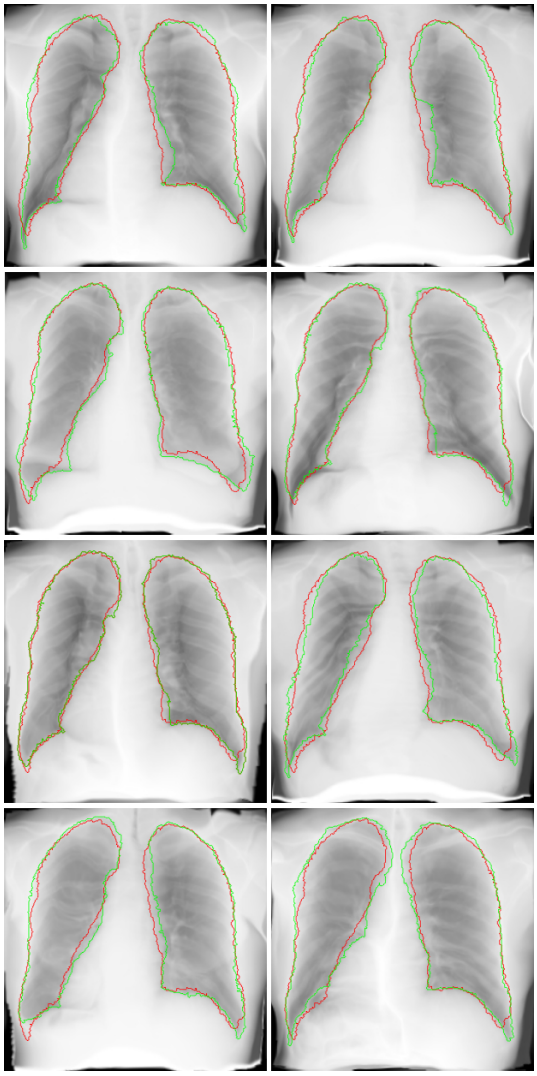


Figure 11: Comparison of the estimated segmentation (red color) and the ground truth (green color) for eight images.

Acknowledgments

This work was partially supported by Association Française contre les Myopathies (AFM: `http://www.afm-france.org`). The authors would also like to thank J-F Deux, G. Bassez, A. Rahmouni from CHU-Henri Mondor for the acquistion of the medical data and the manual

delineation of the region muscles. Last but not least, the authors would like to thank Georg Langs and Ben Glocker for fruitful discussions and advices on the problem.

References

[1] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through mrfs and efficient linear programming," *Medical Image Analysis*, 2008.

[2] Y. Zhan, D. Shen, J. Zeng, L. Sun, G. Fichtinger, J. Moul, and C. Davatzikos, "Targeted Prostate Biopsy Using Statistical Image Analysis," *IEEE T-MI*, 2007.

[3] P. Thompson and A. Toga, "The role of image registration in brain mapping," *Image and Vision Computing*, vol. 19, pp. 3–24, 2001.

[4] H. Park, P.H. Bland, A.O. Hero III, and C.R. Meyer, "Least biased target selection in probabilistic atlas construction," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, 2005, pp. 419–426.

[5] S. Marsland, C.J. Twining, and C.J. Taylor, "Groupwise non-rigid registration using polyharmonic clamped-plate splines," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2003*, 2003, vol. 2, pp. 771–779.

[6] D. Seghers, E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "Construction of a brain template from mr images using state-of-the-art registration and segmentation techniques," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, 2004, vol. 1, pp. 696–703.

[7] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *Neuroimage*, vol. 23, pp. 151–160, 2004.

[8] K.K. Bhatia, J.V. Hajnal, B.K. Puri, A.D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *IEEE ISBI*, 2004.

[9] C. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, and C. Taylor, "A unified information-theoretic approach to groupwise non-rigid registration and model building," in *Internationl Proceedings in Medical Imaging*, 2005, pp. 1–14.

[10] G. Langs, R. Donner, P. Peloschek, and H. Bischof, "Robust autonomus model learning from 2d and 3d data sets," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 2007.

[11] E.G. Miller, N.E. Matsakis, and P.A. Viola, "Learning from one example through shared densities on transforms," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2000, pp. 464–471.

[12] L. Zollei, E. Learned-Miller, E. Grimson, and W. Wells, "Efficient population registration of 3d data," in *IEEE ICCV*, 2005.

[13] S.K. Balci, P. Golland, M. Shenton, and W.M. Wells, "Free-form b-spline deformation model for groupwise registration," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 2007.

[14] B. Ma, H. Park, A. O. Hero, P. H. Bland, and C. Meyer, "Comparing pairwise and simultaneous joint registrations of decorrelating interval exams using entropic graphs," in *IPMI*, 2007, pp. 270–282.

[15] K.H Huebner, *The Finite Element Method for Engineers*, John Willey and Sons, 1975.

[16] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, "Non-rigid registration using free-form deformations: application to breast mr images," *IEEE T-MI*, vol. 18, no. 8, pp. 712–721, Aug. 1999.

[17] Olga Veksler, "Efficient graph-based energy minimization methods in computer vision," 1999.

[18] Vladimir Kolmogorov and Ramin Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 65–81, 2002.

[19] N. Komodakis, G. Tziritas, and N. Paragios, "Fast, approximately optimal solutions for single and dynamic MRFs," in *IEEE CVPR*, 2007.

[20] D. Rueckert, P. Aljabar, R. Heckemann, J. Hajnal, and A. Hammers, "Diffeomorphic registration using b-splines," in *MICCAI*, 2006.

[21] Y. Choi and S. Lee, "Injectivity conditions of 2d and 3d uniform cubic b-spline functions," *Graphical Models*, 2000.

[22] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *AJR*, 2000.

[23] B. van Ginneken, M.B. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database," *Medical Image Analysis*, 2006.

[24] L. Zollei, M. Shenton, W.M. Wells III, and K. Pohl, "The Impact of Atlas Formation Methods on Atlas-Guided Brain Segmentation," *MICCAI*, 2007.

Unité de recherche INRIA Futurs
Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399